

# Convergence and superconvergence analyses of HDG methods for time fractional diffusion problems

Kassem Mustapha · Maher Nour ·  
Bernardo Cockburn

December 8, 2014

**Abstract** We study the hybridizable discontinuous Galerkin (HDG) method for the spatial discretization of time fractional diffusion models with Caputo derivative of order  $0 < \alpha < 1$ . For each time  $t \in [0, T]$ , the HDG approximations are taken to be piecewise polynomials of degree  $k \geq 0$  on the spatial domain  $\Omega$ , the approximations to the exact solution  $u$  in the  $L_\infty(0, T; L_2(\Omega))$ -norm and to  $\nabla u$  in the  $L_\infty(0, T; \mathbf{L}_2(\Omega))$ -norm are proven to converge with the rate  $h^{k+1}$  provided that  $u$  is sufficiently regular, where  $h$  is the maximum diameter of the elements of the mesh. Moreover, for  $k \geq 1$ , we obtain a superconvergence result which allows us to compute, in an elementwise manner, a new approximation for  $u$  converging with a rate  $h^{k+2}$  (ignoring the logarithmic factor), for quasi-uniform spatial meshes. Numerical experiments validating the theoretical results are displayed.

**Keywords** Anomalous diffusion · Time fractional · Discontinuous Galerkin methods · Hybridization · Convergence analysis

## 1 Introduction

In this paper, we study the method resulting after using exact integration in time and a hybridizable discontinuous Galerkin (HDG) method for the spatial discretization of the following time fractional diffusion model problem:

$${}^c\mathcal{D}^{1-\alpha}u(x, t) - \Delta u(x, t) = f(x, t) \quad \text{for } (x, t) \in \Omega \times (0, T], \quad (1a)$$

$$u(x, t) = g(x) \quad \text{for } (x, t) \in \partial\Omega \times (0, T], \quad (1b)$$

---

Support of the King Fahd University of Petroleum and Minerals (KFUPM) through the project FT131011 is gratefully acknowledged.

K. Mustapha E-mail: [kassem@kfupm.edu.sa](mailto:kassem@kfupm.edu.sa) · M. Nour E-mail: [mnoor@kfupm.edu.sa](mailto:mnoor@kfupm.edu.sa)  
Department of Mathematics and Statistics, KFUPM, Saudi Arabia

B. Cockburn E-mail: [cockburn@math.umn.edu](mailto:cockburn@math.umn.edu)  
School of Mathematics, University of Minnesota, USA

with  $u(x, 0) = u_0(x)$  for  $x \in \Omega$ , where  $\Omega$  is a convex polyhedral domain of  $\mathbb{R}^d$  ( $d = 1, 2, 3$ ) with boundary  $\partial\Omega$ ,  $f$ ,  $g$  and  $u_0$  are given functions assumed to be sufficiently regular such that the solution  $u$  of (1) is in the space  $W^{1,1}(0, T; H^2(\Omega))$ , see the regularity analysis in [18] (further regularity assumptions will be imposed later), and  $T > 0$  is a fixed but arbitrary value. Here,  ${}^cD^{1-\alpha}$  denotes time fractional Caputo derivative of order  $\alpha$  defined by

$${}^cD^{1-\alpha}v(t) := \mathcal{I}^\alpha v'(t) := \int_0^t \omega_\alpha(t-s)v'(s)ds \quad \text{with } 0 < \alpha < 1, \quad (2)$$

where  $v'$  denotes the time derivative of the function  $v$  and  $\mathcal{I}^\alpha$  is the Riemann–Liouville (time) fractional integral operator; with  $\omega_\alpha(t) := \frac{t^{\alpha-1}}{\Gamma(\alpha)}$  and  $\Gamma$  being the gamma function.

In this work, we investigate a high-order accurate numerical method for the space discretization for problem (1). Using exact integration in time, we propose to deal with the accuracy issue by developing a high-order HDG method that allows for locally varying spatial meshes and approximation orders which are beneficial to handle problems with low regularity. The HDG methods were introduced in [4] in the framework of steady-state diffusion which share with the classical (hybridized version of the) mixed finite element methods their remarkable convergence and superconvergence properties, [7], as well as the way in which they can be efficiently implemented, [15]. They provide approximations that are more accurate than the ones given by any other DG method for second-order elliptic problems [30]. In [6], a similar method was studied for the fractional subdiffusion problem:

$$u'(x, t) - D^{1-\alpha}\Delta u(x, t) = f(x, t) \quad \text{for } (x, t) \in \Omega \times (0, T], \quad (3)$$

where  $D^{1-\alpha}$  is the Riemann–Liouville fractional time derivative operator,

$$D^{1-\alpha}v(t) := (\mathcal{I}^\alpha v(t))' = \frac{\partial}{\partial t} \int_0^t \omega_\alpha(t-s)v(s)ds. \quad (4)$$

(For other numerical methods of (3), see [2, 8, 9, 20, 23, 24, 27, 28, 35] and related references therein.) When  $f \equiv 0$  (that is, homogeneous case), the two representations (1a) and (3) are different ways of writing the same equation, as they are equivalent under reasonable assumptions on the initial data. However, the numerical methods obtained for each representation are formally different. In [6], the authors extended the approach of the error analysis used in [1] for the heat equation by using several important properties of  $D^{1-\alpha}$ . Indeed, a duality argument was applied (where delicate regularity estimates were required) to prove the superconvergence properties of the method.

We start our work by introducing the spatial semi-discrete HDG method for the model problem (1) in the next section. For sake of implementing the HDG scheme, we discretize in time using a generalized Crank–Nicolson scheme [25]. The existence and uniqueness of the obtained fully discrete scheme will be shown. In Section 3, we prove the main optimal convergence results of the HDG method. Indeed, for each time  $t \in [0, T]$ , we prove that the error of the

HDG approximation to the solution  $u$  of (1) in the  $L_\infty(0, T; L_2(\Omega))$ -norm and to the flux  $\mathbf{q} := -\nabla u$  in the  $L_\infty(0, T; \mathbf{L}_2(\Omega))$ -norm converge with order  $h^{k+1}$  where  $k$  is the polynomial degree and  $h$  is the maximum diameter of the elements of the spatial mesh; see Theorem 2. Some important properties of the fractional integral operator  $\mathcal{I}^\alpha$  are used in our a priori error analysis. In Section 4, for quasi-uniform meshes and whenever  $k \geq 1$ , by a simple element-wise postprocessing with a computation cost that is negligible in comparison with that of obtaining the HDG approximate solution, we obtain a better approximation to  $u$  converging in the  $L_\infty(0, T; L_2(\Omega))$ -norm with a rate of order  $\sqrt{\log(T/h^{2/(\alpha+1)})}h^{k+2}$ ; see Theorem 3. Here, we partially rely on the superconvergence analysis of the postprocessed HDG scheme in [6, Section 5]. In Section 5, we present some numerical tests which indicate the validity of our theoretical optimal convergence rates of the HDG scheme as well as the superconvergence rates of the postprocessed HDG scheme.

Here is a brief history of the numerical methods for problem (1) in the existing literature. For the *one dimensional case*, a box-type scheme based on combining order reduction approach and an  $L_1$ -discretization was considered in [38]. An explicit finite difference (FD) method, we refer the reader to [31]. For an implicit FD scheme in time and Legendre spectral methods in space were studied in [17]. An extension of this work was considered in [16], where a time-space spectral method has been proposed and analyzed. An implicit Crank–Nicolson had been considered in [32] where the stability of the proposed scheme was proven. Two finite difference/element approaches were developed in [36]. Therein, the time direction was approximated by the fractional linear multistep method and the space direction was approximated by the standard finite element method (FEM). A compact difference scheme (fourth order in space) was proposed in [39] for solving problem (1) but with a variable diffusion parameter. The unconditional stability and the global convergence of the scheme were shown. In [34], a high-order local DG (LDG) method for space discretization was studied. Optimal convergence rates was proved.

For the *two- (or three-) dimensional cases*, a standard second-order central difference approximation was used in space, and, for the time stepping, two alternating direction implicit schemes ( $L_1$ -approximation and backward Euler method) were investigated in [37]. A fractional alternating direction implicit scheme for problem (1) in 3D was proposed in [3]. Unique solvability, unconditional stability and convergence in  $H^1$ -norm were shown. A compact fourth order FD method (in space) with operator-splitting techniques was considered in [10]. The Caputo derivative was evaluated by the  $L_1$  approximation, and the second order spatial derivatives were approximated by the fourth-order, compact (implicit) finite differences. In [13], the authors developed two simple fully discrete schemes based on piecewise linear Galerkin FEMs in space and implicit backward differences for the time discretizations. Finally, a high-order accurate (variable) time-stepping discontinuous Petrov-Galerkin that allows low regularity combined with standard finite elements in space was investigated recently in [25]. Stability and error analysis were rigourously studied.

## 2 The HDG method

This section is devoted to defining a scalar approximation  $u_h(t)$  to  $u(t)$ , a vector approximation  $\mathbf{q}_h(t)$  to the flux  $\mathbf{q}(t)$ , and a scalar approximation  $\widehat{u}_h(t)$  to the trace of  $u(t)$  on element boundaries for each time  $t \in [0, T]$ , using a spatial HDG method. We begin by discretizing the domain  $\Omega$  by a conforming triangulation (for simplicity)  $\mathcal{T}_h$  made of simplexes  $K$ ; we denote by  $\partial\mathcal{T}_h$  the set of all the boundaries  $\partial K$  of the elements  $K$  of  $\mathcal{T}_h$ . We denote by  $\mathcal{E}_h$  the union of faces  $F$  of the simplexes  $K$  of the triangulation  $\mathcal{T}_h$ .

Next, we introduce the discontinuous finite element spaces:

$$W_h = \{w \in L^2(\Omega) : w|_K \in \mathcal{P}_k(K) \quad \forall K \in \mathcal{T}_h\}, \quad (5a)$$

$$\mathbf{V}_h = \{\mathbf{v} \in [L_2(\Omega)]^d : \mathbf{v}|_K \in [\mathcal{P}_k(K)]^d \quad \forall K \in \mathcal{T}_h\}, \quad (5b)$$

$$M_h = \{\mu \in L^2(\mathcal{E}_h) : \mu|_F \in \mathcal{P}_k(F) \quad \forall F \in \mathcal{E}_h\}, \quad (5c)$$

where  $\mathcal{P}_k(K)$  is the space of polynomials of total degree at most  $k$  in the spatial variable.

To describe our scheme, we rewrite (1a) as a first order system as follows:  $\mathbf{q} + \nabla u = 0$  and  ${}^cD^{1-\alpha}u + \nabla \cdot \mathbf{q} = f$ . So,  $\mathbf{q}$  and  $u$  satisfy: for  $t \in (0, T]$ ,

$$(\mathbf{q}, \phi) - (u, \nabla \cdot \phi) + \langle u, \phi \cdot \mathbf{n} \rangle = 0 \quad \forall \phi \in H(\text{div}, \Omega), \quad (6a)$$

$$({}^cD^{1-\alpha}u, \chi) - (\mathbf{q}, \nabla \chi) + \langle \mathbf{q} \cdot \mathbf{n}, \chi \rangle = (f, \chi) \quad \forall \chi \in H^1(\Omega). \quad (6b)$$

where  $(v, w) := \sum_{K \in \mathcal{T}_h} (v, w)_K$  and  $\langle v, w \rangle := \sum_{K \in \mathcal{T}_h} \langle v, w \rangle_{\partial K}$ . Throughout the paper, for any domain  $D$  in  $\mathbb{R}^d$ , by  $(u, v)_D = \int_D uv \, dx$  we denote the  $L_2$ -inner product on  $D$ . However, we use instead  $\langle u, v \rangle_D$  for the  $L_2$ -inner product when  $D$  is a domain of  $\mathbb{R}^{d-1}$ . We use  $\|\cdot\|_D$  to denote the  $L^2(D)$ -norm where we drop  $D$  when  $D = \Omega$ . For vector functions  $\mathbf{v}$  and  $\mathbf{w}$ , the notation is similarly defined with the integrand being the dot product  $\mathbf{v} \cdot \mathbf{w}$ . For later use, the norm and semi-norm on any Sobolev space  $X$  are denoted by  $\|\cdot\|_X$  and  $|\cdot|_X$ , respectively. We also denote  $\|\cdot\|_{X(0,T;Y(\Omega))}$  by  $\|\cdot\|_{X(Y)}$ .

For each  $t > 0$ , the HDG method provides approximations  $u_h(t) \in W_h$ ,  $\mathbf{q}_h(t) \in \mathbf{V}_h$ , and  $\widehat{u}_h(t) \in M_h$  of  $u(t)$ ,  $\mathbf{q}(t)$ , and the trace of  $u(t)$ , respectively. These are determined by requiring that

$$(\mathbf{q}_h, \mathbf{r}) - (u_h, \nabla \cdot \mathbf{r}) + \langle \widehat{u}_h, \mathbf{r} \cdot \mathbf{n} \rangle = 0, \quad \forall \mathbf{r} \in \mathbf{V}_h, \quad (7a)$$

$$({}^cD^{1-\alpha}u_h, w) - (\mathbf{q}_h, \nabla w) + \langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, w \rangle = (f, w), \quad \forall w \in W_h, \quad (7b)$$

$$\langle \widehat{u}_h, \mu \rangle_{\partial\Omega} = \langle g, \mu \rangle_{\partial\Omega}, \quad \forall \mu \in M_h, \quad (7c)$$

$$\langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle - \langle \widehat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle_{\partial\Omega} = 0, \quad \forall \mu \in M_h, \quad (7d)$$

and take the numerical trace for the flux as

$$\widehat{\mathbf{q}}_h = \mathbf{q}_h + \tau(u_h - \widehat{u}_h)\mathbf{n} \quad \text{on } \partial\mathcal{T}_h, \quad (7e)$$

for some nonnegative stabilization function  $\tau$  defined on  $\partial\mathcal{T}_h$ ; we assume that, for each element  $K \in \mathcal{T}_h$ ,  $\tau|_{\partial K}$  is constant on each of its faces. At  $t = 0$ ,  $u_h(0) := \Pi_w u_0$  where the operator  $\Pi_w$  is defined in (9).

The first two equations are inspired in the weak form of the fractional differential equations satisfied by the exact solution, (6). The form of the numerical trace given by (7d) allows us to express  $(u_h, \mathbf{q}_h, \hat{\mathbf{q}}_h)$  elementwise in terms of  $\hat{u}_h$  and  $f$  by using equations (7a), (7b) and (7e). Then, the numerical trace  $\hat{u}_h$  is determined by as the solution of the transmission condition (7d), which enforces the single-valuedness of the normal component of the numerical trace  $\hat{\mathbf{q}}_h$ , and the boundary condition (7c). Thus, the only globally-coupled degrees of freedom are those of  $\hat{u}_h$ .

In our experiments, to implement our spatial semi-discrete HDG scheme (7), we use for simplicity a generalized Crank-Nicolson (CN) scheme for time discretization, see [22, 25]. Formally, the CN scheme is second-order accurate provided that the continuous solution is sufficiently regular. To this end, we introduce a uniform partition of the time interval  $[0, T]$  given by the points:  $t_i = i\delta$  for  $i = 0, \dots, N$ , with  $\delta = T/N$  being the time-step size. We take  $\delta$  to be sufficiently small so that the spatial discretizations errors are dominant.

The time-stepping CN combined with the above HDG method provides approximations  $u_h^j \in W_h$ ,  $\mathbf{q}_h^j \in \mathbf{V}_h$ , and  $\hat{u}_h^j \in M_h$  of  $u(t_j)$ ,  $\mathbf{q}(t_j)$ , and the trace of  $u(t_j)$ , respectively, for  $j = 1, \dots, N$ . Starting from  $u_h^0 = \Pi_W u_0$ , and with appropriate choices of  $\mathbf{q}_h^0$  and  $\hat{u}_h^0$ , our fully discrete scheme is defined by:

$$\begin{aligned} (\mathbf{q}_h^{j-\frac{1}{2}}, \mathbf{r}) - (u_h^{j-\frac{1}{2}}, \nabla \cdot \mathbf{r}) + \langle \hat{u}_h^{j-\frac{1}{2}}, \mathbf{r} \cdot \mathbf{n} \rangle &= 0, & \forall \mathbf{r} \in \mathbf{V}_h, \\ (\mathcal{J}_\alpha \bar{u}_h(t_j), w) - (\mathbf{q}_h^{j-\frac{1}{2}}, \nabla w) + \langle \hat{\mathbf{q}}_h^{j-\frac{1}{2}}, \mathbf{n}, w \rangle &= (f^{j-\frac{1}{2}}, w), & \forall w \in W_h, \\ \langle \hat{u}_h^j, \mu \rangle_{\partial\Omega} &= \langle g, \mu \rangle_{\partial\Omega}, & \forall \mu \in M_h, \\ \langle \hat{\mathbf{q}}_h^j \cdot \mathbf{n}, \mu_1 \rangle - \langle \hat{\mathbf{q}}_h^j \cdot \mathbf{n}, \mu_1 \rangle_{\partial\Omega} &= 0, & \forall \mu_1 \in M_h, \end{aligned} \quad (8)$$

where  $f^{j-\frac{1}{2}} := \frac{1}{2}(f(t_{j-1}) + f(t_j))$ ,  $\hat{\mathbf{q}}_h^j = \mathbf{q}_h^j + \tau(u_h^j - \hat{u}_h^j)\mathbf{n}$  on  $\partial\mathcal{T}_h$ ,

$$\mathcal{J}_\alpha \bar{u}_h(t_j) = \int_{t_{j-1}}^{t_j} \int_0^t \omega_\alpha(t-s) \bar{u}_h(s) ds dt,$$

with  $\bar{u}_h(s) := \delta^{-1}(u_h^i - u_h^{i-1})$  for  $s \in (t_{i-1}, t_i)$ ,  $\mathbf{q}_h^{j-\frac{1}{2}} := \frac{1}{2}(\mathbf{q}_h^j + \mathbf{q}_h^{j-1})$ , and the functions  $u_h^{j-\frac{1}{2}}$ ,  $\hat{u}_h^{j-\frac{1}{2}}$ , and  $\hat{\mathbf{q}}_h^{j-\frac{1}{2}}$  are similarly defined.

For each  $1 \leq j \leq N$ , (8) amounts to a square linear system. Thus the existence of the CN HDG solution follows from its uniqueness. We prove the uniqueness by induction hypothesis on  $j$ . We let  $f^{i-\frac{1}{2}}$  (for  $1 \leq i \leq j$ ) and  $g$  be identically zero in (8), we assume that  $(u_h^i, \mathbf{q}_h^i, \hat{u}_h^i) \equiv (0, \mathbf{0}, 0)$  for  $1 \leq i \leq j-1$  and the task is to show that this holds true for  $i = j$ . To do so, choose  $\mathbf{r} = \mathbf{q}_h^j$ ,  $w = u_h^j$ ,  $\mu = \hat{\mathbf{q}}_h^j \cdot \mathbf{n}$  and  $\mu_1 = \hat{u}_h^j$  in (8) and then simplify, yield

$$\begin{aligned} \|\mathbf{q}_h^j\|^2 - (u_h^j, \nabla \cdot \mathbf{q}_h^j) + \langle \hat{u}_h^j, \mathbf{q}_h^j \cdot \mathbf{n} \rangle &= 0, \\ (\mathcal{J}_\alpha \bar{u}_h(t_j), u_h^j) - (\mathbf{q}_h^j, \nabla u_h^j) + \langle \hat{\mathbf{q}}_h^j \cdot \mathbf{n}, u_h^j \rangle &= 0, \\ \langle \hat{\mathbf{q}}_h^j \cdot \mathbf{n}, \hat{u}_h^j \rangle &= 0, \end{aligned}$$

Since  $(u_h^j, \nabla \cdot \mathbf{q}_h^j) = \langle u_h^j, \mathbf{q}_h^j \cdot \mathbf{n} \rangle - (\mathbf{q}_h^j, \nabla u_h^j)$ , adding the above equations give

$$(\mathcal{I}^\alpha \bar{u}_h(t_j), u_h^j) + \|\mathbf{q}_h^j\|^2 + \langle \hat{u}_h^j - u_h^j, (\mathbf{q}_h^j - \hat{\mathbf{q}}_h^j) \cdot \mathbf{n} \rangle = 0.$$

Hence, by the induction hypothesis and the identity  $(\mathbf{q}_h^j - \hat{\mathbf{q}}_h^j) \cdot \mathbf{n} = \tau(u_h^j - \hat{u}_h^j)$  on  $\partial\mathcal{T}_h$ , we notes that

$$\int_0^{t_j} (\mathcal{I}^\alpha \bar{u}_h(t), \bar{u}_h(t)) dt + \|\mathbf{q}_h^j\|^2 + \|\sqrt{\tau}(\hat{u}_h^j - u_h^j)\|_{\partial\mathcal{T}_h}^2 = 0,$$

and therefore, the use of the coercivity property of  $\mathcal{I}^\alpha$  (see (18)) completes the proof.

### 3 Error estimates

In this section, we carry our a priori error analysis of the HDG method. Following [1, 5, 6], we start by defining the projections below which play the comparison function role in the error analysis.

For each  $t \in (0, T]$ , we assume that  $\mathbf{q}(t) \in [H^1(\mathcal{T}_h)]^d$  and  $u(t) \in H^1(\mathcal{T}_h)$ , where  $H^1(\mathcal{T}_h) = \prod_{K \in \mathcal{T}_h} H^1(K)$ , the projections  $\Pi_V \mathbf{q}(t) \in \mathbf{V}_h$  and  $\Pi_W u(t) \in W_h$  are defined by: on each simplex  $K \in \mathcal{T}_h$  and for all faces  $F$  of  $K$ ,

$$(\Pi_V \mathbf{q}(t), \mathbf{v})_K = (\mathbf{q}(t), \mathbf{v})_K, \quad (9a)$$

$$(\Pi_W u(t), w)_K = (u(t), w)_K, \quad (9b)$$

$$\langle \Pi_V \mathbf{q}(t) \cdot \mathbf{n} + \tau \Pi_W u(t), \mu \rangle_F = \langle \mathbf{q}(t) \cdot \mathbf{n} + \tau u(t), \mu \rangle_F, \quad (9c)$$

for al  $\mathbf{v} \in [\mathcal{P}_{k-1}(K)]^d$ ,  $w \in \mathcal{P}_{k-1}(K)$  and  $\mu \in \mathcal{P}_k(F)$ . This projection introduced in [5] to study HDG methods for the steady-state diffusion problem and also used in the error analyses of HDG methods for classical diffusion [1] as well as for fractional subdiffusion [6] problems. Its approximation properties are described in the following result.

**Theorem 1** ([5]) *Suppose  $\tau|_{\partial K}$  is nonnegative and  $\tau_K^{\max} := \max \tau|_{\partial K} > 0$ . Then the system (9) is uniquely solvable for  $\Pi_V \mathbf{q}$  and  $\Pi_W u$ . Furthermore, there is a constant  $C$  independent of  $K$  and  $\tau$  such that for each  $t \in (0, T]$ ,*

$$\begin{aligned} \|e_q(t)\|_K &\leq C h_K^{k+1} (|\mathbf{q}(t)|_{\mathbf{H}^{k+1}(K)} + \tau_K^* |u(t)|_{H^{k+1}(K)}), \\ \|e_u(t)\|_K &\leq C h_K^{k+1} (|u(t)|_{H^{k+1}(K)} + |\nabla \cdot \mathbf{q}(t)|_{H^k(K)} / \tau_K^{\max}) \end{aligned}$$

where  $e_q := \Pi_V \mathbf{q} - \mathbf{q}$  and  $e_u := \Pi_W u - u$ . Here  $\tau_K^* := \max \tau|_{\partial K \setminus F^*}$ , where  $F^*$  is a face of  $K$  at which  $\tau|_{\partial K}$  is maximum.

Note that the approximation error of the projection is of order  $k+1$  provided that the stabilization function is such that both  $\tau_K^*$  and  $1/\tau_K^{\max}$  are uniformly bounded and the exact solution is sufficiently regular.

Thus, the main task now is to estimate the terms  $\varepsilon_u := \Pi_W u - u_h$  and  $\varepsilon_q := \Pi_V \mathbf{q} - \mathbf{q}_h$ . For convenience, we further introduce the following notations:

$\varepsilon_{\hat{u}} := P_M u - \hat{u}_h$  and  $\varepsilon_{\hat{q}} := \mathbf{P}_M \mathbf{q} - \hat{\mathbf{q}}_h$  where  $P_M$  denotes the  $L^2$ -orthogonal projection onto  $M_h$ , and  $\mathbf{P}_M$  denotes the vector-valued projection each of whose components are equal to  $P_M$ .

The projection of the errors satisfy the equations stated in the next lemma.

**Lemma 1** *For each  $t > 0$ , we have*

$$(\varepsilon_{\mathbf{q}}, \mathbf{r}) - (\varepsilon_u, \nabla \cdot \mathbf{r}) + \langle \varepsilon_{\hat{u}}, \mathbf{r} \cdot \mathbf{n} \rangle = (e_{\mathbf{q}}, \mathbf{r}), \quad \forall \mathbf{r} \in \mathbf{V}_h \quad (11a)$$

$$(\mathcal{I}^\alpha \varepsilon'_u, w) - (\varepsilon_{\mathbf{q}}, \nabla w) + \langle \varepsilon_{\hat{q}}, \mathbf{n}, w \rangle = (\mathcal{I}^\alpha e'_u, w), \quad \forall w \in W_h \quad (11b)$$

$$\langle \varepsilon_{\hat{u}}, \mu \rangle_{\partial\Omega} = 0, \quad \forall \mu \in M_h \quad (11c)$$

$$\langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \mu \rangle - \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \mu \rangle_{\partial\Omega} = 0, \quad \forall \mu \in M_h \quad (11d)$$

where

$$\varepsilon_{\hat{q}} \cdot \mathbf{n} := \varepsilon_{\mathbf{q}} \cdot \mathbf{n} + \tau(\varepsilon_u - \varepsilon_{\hat{u}}) \quad \text{on } \partial\mathcal{T}_h. \quad (11e)$$

*Proof* From (6), we recall that  $\mathbf{q}$  and  $u$  satisfy the equations

$$\begin{aligned} (\mathbf{q}, \mathbf{r}) - (u, \nabla \cdot \mathbf{r}) + \langle u, \mathbf{r} \cdot \mathbf{n} \rangle &= 0 \quad \forall \mathbf{r} \in \mathbf{V}_h, \\ (\mathcal{I}^\alpha u', w) - (\mathbf{q}, \nabla w) + \langle \mathbf{q} \cdot \mathbf{n}, w \rangle &= (f, w) \quad \forall w \in W_h. \end{aligned}$$

By (9a), (9b) and the fact that  $P_M$  is the  $L^2$ -projection into  $M_h$ , we get

$$(\mathbf{I}_V \mathbf{q}, \mathbf{r}) - (\Pi_W u, \nabla \cdot \mathbf{r}) + \langle P_M u, \mathbf{r} \cdot \mathbf{n} \rangle = (e_{\mathbf{q}}, \mathbf{r}), \quad \forall \mathbf{r} \in \mathbf{V}_h \quad (13)$$

$$\begin{aligned} (\mathcal{I}^\alpha (\Pi_W u)', w) - (\mathbf{I}_V \mathbf{q}, \nabla w) + \langle \mathbf{I}_V \mathbf{q} \cdot \mathbf{n} + \tau(\Pi_W u - P_M u), w \rangle \\ = (f + \mathcal{I}^\alpha e'_u, w), \quad \forall w \in W_h, \end{aligned} \quad (14)$$

given that, for each element  $K \in \mathcal{T}_h$ ,  $\tau$  is constant on each face  $F$  of  $K$ .

Subtracting the equations (7a) and (7b) from (13) and (14), respectively, we obtain equations (11a) and (11b), respectively. The equation (11c) follows directly from the equation (7c) and (1b)

By the definition of  $\varepsilon_{\hat{q}}$  and since  $P_M$  is the  $L^2$ -projection into  $M_h$ , we have

$$\begin{aligned} \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \mu \rangle - \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \mu \rangle_{\partial\Omega} &= \langle (p_M \mathbf{q} - \hat{\mathbf{q}}_h) \cdot \mathbf{n}, \mu \rangle - \langle (p_M \mathbf{q} - \hat{\mathbf{q}}_h) \cdot \mathbf{n}, \mu \rangle_{\partial\Omega} \\ &= \langle (\mathbf{q} - \hat{\mathbf{q}}_h) \cdot \mathbf{n}, \mu \rangle - \langle (\mathbf{q} - \hat{\mathbf{q}}_h) \cdot \mathbf{n}, \mu \rangle_{\partial\Omega} \\ &= [\langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle - \langle \mathbf{q} \cdot \mathbf{n}, \mu \rangle_{\partial\Omega}] - [\langle \hat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle - \langle \hat{\mathbf{q}}_h \cdot \mathbf{n}, \mu \rangle_{\partial\Omega}] = 0 \end{aligned}$$

where in the last equality we used that  $\mathbf{q}$  is in  $\mathbf{H}(\text{div}, \Omega)$  and equation (7d). Thus, the identity (11d) holds. For the proof of (11e),

$$\begin{aligned} \varepsilon_{\hat{q}} \cdot \mathbf{n} &= P_M(\mathbf{q} \cdot \mathbf{n}) - (\mathbf{q}_h \cdot \mathbf{n} + \tau(u_h - \hat{u}_h)) && \text{by (7e),} \\ &= (\mathbf{I}_V \mathbf{q} \cdot \mathbf{n} + \tau(\Pi_W u - P_M u)) - (\mathbf{q}_h \cdot \mathbf{n} + \tau(u_h - \hat{u}_h)) && \text{by (9c),} \\ &= \varepsilon_{\mathbf{q}} \cdot \mathbf{n} + \tau(\varepsilon_u - \varepsilon_{\hat{u}}). \quad \square \end{aligned}$$

**Lemma 2** Let  $S_h := \|\sqrt{\tau}(\varepsilon_u - \varepsilon_{\hat{u}})\|_{\partial\mathcal{T}_h}$  and  $c_\alpha := \cos(\frac{\alpha\pi}{2})$ . For  $T > 0$ ,

$$\begin{aligned} \int_0^T (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) dt + \|\varepsilon_{\mathbf{q}}(T)\|^2 + S_h^2(T) &\leq \|\varepsilon_{\mathbf{q}}(0)\|^2 + S_h^2(0) \\ &+ \frac{1}{c_\alpha^2} \int_0^T (\mathcal{I}^\alpha e'_u, e'_u) dt + 2 \int_0^T (e'_{\mathbf{q}}, \varepsilon_{\mathbf{q}}) dt. \end{aligned}$$

*Proof* Since  $(\varepsilon_u, \nabla \cdot \mathbf{r}) = -(\nabla \varepsilon_u, \mathbf{r}) + \langle \varepsilon_u, \mathbf{r} \cdot \mathbf{n} \rangle$ , (11a) can be rewritten as:

$$(\varepsilon_{\mathbf{q}}, \mathbf{r}) + (\nabla \varepsilon_u, \mathbf{r}) + \langle \varepsilon_{\hat{u}} - \varepsilon_u, \mathbf{r} \cdot \mathbf{n} \rangle = (e_{\mathbf{q}}, \mathbf{r}).$$

A time differentiation of both sides yields,

$$(\varepsilon'_{\mathbf{q}}, \mathbf{r}) + (\nabla \varepsilon'_u, \mathbf{r}) + \langle \varepsilon'_{\hat{u}} - \varepsilon'_u, \mathbf{r} \cdot \mathbf{n} \rangle = (e'_{\mathbf{q}}, \mathbf{r}).$$

Setting  $\mathbf{r} = \varepsilon_{\mathbf{q}}$  and choose  $w = \varepsilon'_u$  in equation (11b), we observe,

$$\begin{aligned} (\varepsilon'_{\mathbf{q}}, \varepsilon_{\mathbf{q}}) + (\nabla \varepsilon'_u, \varepsilon_{\mathbf{q}}) + \langle \varepsilon'_{\hat{u}} - \varepsilon'_u, \varepsilon_{\mathbf{q}} \cdot \mathbf{n} \rangle &= (e'_{\mathbf{q}}, \varepsilon_{\mathbf{q}}), \\ (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) - (\varepsilon_{\mathbf{q}}, \nabla \varepsilon'_u) + \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \varepsilon'_u \rangle &= (\mathcal{I}^\alpha e'_u, \varepsilon'_u). \end{aligned}$$

Combine the above two equations and using  $(\varepsilon'_{\mathbf{q}}, \varepsilon_{\mathbf{q}}) = \frac{1}{2} \frac{d}{dt} \|\varepsilon_{\mathbf{q}}\|^2$ , we obtain

$$(\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) + \frac{1}{2} \frac{d}{dt} \|\varepsilon_{\mathbf{q}}\|^2 + \psi_h = (\mathcal{I}^\alpha e'_u, \varepsilon'_u) + (e'_{\mathbf{q}}, \varepsilon_{\mathbf{q}}). \quad (15)$$

where

$$\psi_h = \langle \varepsilon'_{\hat{u}} - \varepsilon'_u, \varepsilon_{\mathbf{q}} \cdot \mathbf{n} \rangle + \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \varepsilon'_u \rangle.$$

A time differentiation of (11c) follows by choosing  $\mu = \varepsilon_{\hat{q}} \cdot \mathbf{n}$  and then using (11d) yield  $\langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \varepsilon'_{\hat{u}} \rangle_{\partial\Omega} = \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \varepsilon'_{\hat{u}} \rangle = 0$ . Thus, by (11e),

$$\psi_h = \langle \varepsilon'_{\hat{u}} - \varepsilon'_u, (\varepsilon_{\mathbf{q}} - \varepsilon_{\hat{q}}) \cdot \mathbf{n} \rangle = \langle \tau(\varepsilon'_u - \varepsilon'_{\hat{u}}), (\varepsilon_u - \varepsilon_{\hat{u}}) \rangle = \frac{1}{2} \frac{d}{dt} S_h^2(t). \quad (16)$$

Now, integrating (15) over the time interval  $[0, T]$  and using (16), we get

$$\int_0^T (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) dt + \frac{1}{2} \int_0^T \frac{d}{dt} [\|\varepsilon_{\mathbf{q}}\|^2 + S_h^2] dt = \int_0^T (\mathcal{I}^\alpha e'_u, \varepsilon'_u) dt + \int_0^T (e'_{\mathbf{q}}, \varepsilon_{\mathbf{q}}) dt.$$

Therefore,

$$\begin{aligned} 2 \int_0^T (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) dt + \|\varepsilon_{\mathbf{q}}(T)\|^2 + S_h^2(T) \\ = \|\varepsilon_{\mathbf{q}}(0)\|^2 + S_h^2(0) + 2 \int_0^T (\mathcal{I}^\alpha e'_u, \varepsilon'_u) dt + 2 \int_0^T (e'_{\mathbf{q}}, \varepsilon_{\mathbf{q}}) dt. \end{aligned} \quad (17)$$

An application of the continuity property of the fractional derivative operator  $\mathcal{I}^\alpha$  (see [29, Lemma 3.6]), yields

$$2 \left| \int_0^T (\mathcal{I}^\alpha e'_u, \varepsilon'_u) dt \right| \leq \frac{1}{c_\alpha^2} \int_0^T (\mathcal{I}^\alpha e'_u, e'_u) dt + \int_0^T (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) dt.$$

Finally, inserting this in (17) and simplifying will complete the proof.



To be ready to show the main error bounds of the HDG method, we recall from [29, Lemma 3.6] the following coercivity property of the operator  $\mathcal{I}^\alpha$ . For any  $v \in \mathcal{C}(0, T; L_2(\Omega))$ , we have

$$\text{Coercivity property : } \int_0^T (\mathcal{I}^\alpha v(t), v(t)) dt \geq c_\alpha \int_0^T \|\mathcal{I}^{\frac{\alpha}{2}} v(t)\|^2 dt. \quad (18)$$

**Theorem 2** Assume that  $u \in \mathcal{C}^1(0, T; H^{k+1}(\Omega))$  and  $\mathbf{q} \in \mathcal{C}^1(0, T; \mathbf{H}^{k+1}(\Omega))$ . Assume also that  $\tau_K^*$  and  $1/\tau_K^{\max}$  are bounded by  $\mathbf{C}$ . Then we have that

$$\|(u - u_h)(T)\| + \|(\mathbf{q} - \mathbf{q}_h)(T)\| \leq C_1(1 + T) h^{k+1}.$$

The constant  $C_1$  only depends on  $\mathbf{C}$ ,  $\alpha$ ,  $\|u\|_{\mathcal{C}^1(H^{k+1})}$ , and on  $\|\mathbf{q}\|_{\mathcal{C}^1(H^{k+1})}$ .

*Proof* From the decompositions:  $u - u_h = \varepsilon_u - e_u$  and  $\mathbf{q} - \mathbf{q}_h = \varepsilon_{\mathbf{q}} - e_{\mathbf{q}}$ , and the error projection in Theorem 1, we have

$$\|(u - u_h)(T)\| + \|(\mathbf{q} - \mathbf{q}_h)(T)\| \leq C_1 h^{k+1} + \|\varepsilon_u(T)\| + \|\varepsilon_{\mathbf{q}}(T)\|.$$

The task now is to estimate  $\|\varepsilon_u(T)\|$  and  $\|\varepsilon_{\mathbf{q}}(T)\|$ . From Lemma 2, for  $t \geq 0$ , we have  $E^2(t) \leq A(t) + 2 \int_0^t B(s) E(s) ds$  where

$$\begin{aligned} A(t) &:= \|\varepsilon_{\mathbf{q}}(0)\|^2 + S_h^2(0) + \frac{1}{c_\alpha^2} \int_0^t (\mathcal{I}^\alpha e'_u, e'_u) ds, \quad B(t) := \|e'_{\mathbf{q}}(t)\|, \\ E(t) &:= \left( \|\varepsilon_{\mathbf{q}}(t)\|^2 + \int_0^t (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) ds \right)^{\frac{1}{2}}. \end{aligned}$$

(Note that  $A$  and  $B$  are nonnegative functions.) Thus, an application of the integral inequality (see [6, Lemma 4]) yields

$$E(T) \leq \max_{t \in (0, T)} A^{\frac{1}{2}}(t) + \int_0^T B(s) ds \quad \text{for any } T > 0.$$

Hence,

$$\begin{aligned} \|\varepsilon_{\mathbf{q}}(T)\|^2 + \int_0^T (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) ds &\leq C \left( \|\varepsilon_{\mathbf{q}}(0)\|^2 + S_h^2(0) \right. \\ &\quad \left. + \int_0^T \left( \frac{1}{c_\alpha^2} \|\mathcal{I}^\alpha e'_u\| \|e'_u\| + T \|e'_{\mathbf{q}}\|^2 \right) ds \right). \quad (19) \end{aligned}$$

However, since  $\varepsilon_u(t) = \int_0^t \varepsilon'_u(s) ds = \mathcal{I}^{1-\frac{\alpha}{2}}(\mathcal{I}^{\frac{\alpha}{2}} \varepsilon'_u)(t)$  (because  $\varepsilon_u(0) = 0$ ), by the Cauchy-Schwarz inequality and the coercivity property of the operator  $\mathcal{I}^\alpha$ ,

$$\begin{aligned} \|\varepsilon_u(t)\|^2 &\leq \left( \int_0^t \omega_{1-\frac{\alpha}{2}}(t-s) \|\mathcal{I}^{\frac{\alpha}{2}} \varepsilon'_u(s)\| ds \right)^2 \\ &\leq \int_0^t \omega_{1-\frac{\alpha}{2}}^2(s) ds \int_0^t \|\mathcal{I}^{\frac{\alpha}{2}} \varepsilon'_u(s)\|^2 ds \\ &= \frac{t^{1-\alpha}}{(1-\alpha)\Gamma^2(1-\frac{\alpha}{2})} \int_0^t \|\mathcal{I}^{\frac{\alpha}{2}} \varepsilon'_u(s)\|^2 ds \\ &\leq \frac{t^{1-\alpha}}{(1-\alpha)\Gamma^2(1-\frac{\alpha}{2})c_\alpha} \int_0^t (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon'_u) ds. \end{aligned}$$

Therefore, combining (19) with the above bound, and apply Theorem 1 for the time derivative error projections  $e'_u$  and  $e'_q$ , we obtain

$$\|\varepsilon_q(t)\|^2 + \|\varepsilon_u(t)\|^2 \leq C_1^2(1+T)^2 \left( \|\varepsilon_q(0)\|^2 + S_h^2(0) + h^{2k+2} \right).$$

To complete the proof, we need to bound  $\|\varepsilon_q(0)\|^2 + S_h^2(0)$ .

Since  $(\varepsilon_u, \nabla \cdot \mathbf{r}) = -(\nabla \varepsilon_u, \mathbf{r}) + \langle \varepsilon_u, \mathbf{r} \cdot \mathbf{n} \rangle$ , setting  $\mathbf{r} = \varepsilon_q$  in (11a) and  $w = \varepsilon'_u$  in (11b) yield

$$\begin{aligned} \|\varepsilon_q\|^2 + (\nabla \varepsilon_u, \varepsilon_q) + \langle \varepsilon_u - \varepsilon_{\hat{u}}, \varepsilon_q \cdot \mathbf{n} \rangle &= (e_q, \varepsilon_q), \\ (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon_u) - (\varepsilon_q, \nabla \varepsilon_u) + \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \varepsilon_u \rangle &= (\mathcal{I}^\alpha e'_u, \varepsilon_u). \end{aligned}$$

Adding the above equations, and using  $\langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \varepsilon_{\hat{u}} \rangle = 0$  (this follows by choosing  $\mu = \varepsilon_{\hat{q}} \cdot \mathbf{n}$  in (11c) and  $\mu = \varepsilon_{\hat{u}}$  in (11d)) and (11e), yield

$$(\mathcal{I}^\alpha \varepsilon'_u, \varepsilon_u) + \|\varepsilon_q\|^2 + S_h^2 = (\mathcal{I}^\alpha e'_u, \varepsilon_u) + (e_q, \varepsilon_q).$$

Now, integrating over the time interval  $[0, t]$ , observing that

$$\int_0^t (\mathcal{I}^\alpha \varepsilon'_u, \varepsilon_u) ds = \int_0^t (D^{1-\alpha} \varepsilon_u, \varepsilon_u) ds \geq 0,$$

(the last inequality follows from the nonnegativity property of the Riemann–Liouville fractional derivative operator  $D^{1-\alpha}$ , see [19, Section 2]) and using the inequality  $(e_q, \varepsilon_q) \leq \frac{1}{2}\|e_q\|^2 + \frac{1}{2}\|\varepsilon_q\|^2$ , we get

$$\int_0^t \left[ \frac{1}{2}\|\varepsilon_q\|^2 + S_h^2 \right] ds \leq \int_0^t (\|\mathcal{I}^\alpha e'_u\| \|\varepsilon_u\| + \frac{1}{2}\|e_q\|^2) ds.$$

Therefore, by the mean value theorem for integrals, there exist  $t^*, \tilde{t} \in (0, t)$  such that

$$t \left( \frac{1}{2}\|\varepsilon_q(t^*)\|^2 + S_h^2(t^*) \right) \leq t \left( \|\varepsilon_u(\tilde{t})\| \max_{s \in (0, t)} \|\mathcal{I}^\alpha e'_u(s)\| + \frac{1}{2} \max_{s \in (0, t)} \|e_q(s)\|^2 \right).$$

Finally, simplify by  $t$ , then take  $\lim_{t \downarrow 0}$  to both sides and using  $\varepsilon_u(0) = 0$ , we observe that  $\|\varepsilon_q(0)\|^2 + S_h^2(0) \leq \|e_q(0)\|^2 \leq C_1 h^{2k+2}$  by the error estimate of  $e_q$  given in Theorem 1. The proof is completed now.

#### 4 Superconvergence and post-processing

In this section, we seek a better approximation to  $u$  by means of an element-by-element postprocessing. We begin by describing such approximation, then we show how to get our superconvergence result by a duality argument.

Following [1, 12], for each  $t \in [0, T]$ , we define the postprocessed HDG solution  $u_h^*(t) \in \mathcal{P}_{k+1}(K)$  to  $u(t)$  for each simplex  $K \in \mathcal{T}_h$ , as follows:

$$(u_h^*(t), 1)_K = (u_h(t), 1)_K \quad (20a)$$

$$(\nabla u_h^*(t), \nabla w)_K = -(\mathbf{q}_h(t), \nabla w)_K \quad \forall w \in \mathcal{P}_{k+1}(K). \quad (20b)$$

Let  $P_0$  be the  $L^2(\Omega)$ -projection into the space of functions which are constant on each element  $K \in \mathcal{T}_h$ . One may show that

$$\|(u - u_h^*)(T)\|_K \leq C h_K^{k+2} |u(T)|_{H^{k+2}(K)} + \|P_0 \varepsilon_u(T)\|_K + C h \|\varepsilon_{\mathbf{q}}(T)\|_K. \quad (21)$$

The main task now is to show that the term  $\|P_0 \varepsilon_u\|$  is of order  $O(h^{k+2})$ . Then the postprocessed approximation  $u_h^*$  would converge faster than the original approximation  $u_h$ . Noting that  $\|P_0 \varepsilon_u(T)\| = \sup_{\Theta \in C_0^\infty(\Omega)} \frac{(P_0 \varepsilon_u(T), \Theta)}{\|\Theta\|}$ . To estimate the expression  $(P_0 \varepsilon_u(T), \Theta)$ , we use the traditional duality approach by using the solution of the dual problem

$$\Phi + \nabla \Psi = 0 \quad \text{and} \quad (\mathcal{I}^{\alpha*} \Psi)' - \nabla \cdot \Phi = 0 \quad \text{on } \Omega \times (0, T), \quad (22)$$

with  $\Psi = 0$  on  $\partial\Omega \times (0, T)$  and  $\Psi(T) = \Theta$  on  $\Omega$ , where  $\mathcal{I}^{\alpha*}$  is the adjoint operator of  $\mathcal{I}^\alpha$  defined by [28]:

$$\mathcal{I}^{\alpha*} \psi(t) = \int_t^T \omega_\alpha(s-t) \psi(s) ds.$$

Integrating  $(\mathcal{I}^{\alpha*} \Psi)' - \nabla \cdot \Phi = 0$  over the time interval  $(t, T)$ , we obtain

$$\mathcal{I}^{\alpha*} \Psi(t) + \int_t^T \nabla \cdot \Phi(s) ds = 0. \quad (23)$$

We define now the adjoint  $D^{\alpha*}$  of the Riemann–Liouville fractional derivative operator  $D^\alpha$  (see (4) for the definition of  $D^\alpha$ ) as follows [28]: for  $t \in (0, T)$ ,

$$D^{\alpha*} v(t) = -\frac{\partial}{\partial t} \int_t^T \omega_{1-\alpha}(s-t) v(s) ds \quad \text{for any } v \in \mathcal{C}^1(0, T).$$

Since  $\int_t^q \omega_\alpha(s-t) \omega_{1-\alpha}(q-s) ds = 1$ , it is easy to see that  $\mathcal{I}^{\alpha*}$  is the *right-inverse* of  $D^{\alpha*}$ , that is,  $D^{\alpha*}(\mathcal{I}^{\alpha*} \Psi)(t) = \Psi(t)$ . Hence, using this after applying the operator  $D^{\alpha*}$  to both sides of (23), yields

$$\Psi(t) + D^{\alpha*} \left( \int_t^T \nabla \cdot \Phi(q) dq \right) = 0. \quad (24)$$

However, since

$$\begin{aligned} D^{\alpha*} \left( \int_t^T \nabla \cdot \Phi(s) ds \right) &= -\frac{\partial}{\partial t} \int_t^T \omega_{1-\alpha}(s-t) \int_s^T \nabla \cdot \Phi(q) dq ds \\ &= -\frac{\partial}{\partial t} \int_t^T \nabla \cdot \Phi(q) \int_t^q \omega_{1-\alpha}(s-t) ds dq \\ &= -\frac{\partial}{\partial t} \int_t^T \omega_{2-\alpha}(q-t) \nabla \cdot \Phi(q) dq \\ &= \int_t^T \omega_{1-\alpha}(s-t) \nabla \cdot \Phi(s) ds, \end{aligned}$$

differentiating both sides of (24) with respect to  $t$ , yield  $\Psi' - \nabla \cdot \mathbf{D}^{\alpha*} \Phi = 0$ . Therefore, an alternative formulation of the dual problem (22) is given by:

$$\Phi + \nabla \Psi = 0 \quad \text{on } \Omega \times (0, T), \quad (25a)$$

$$\Psi' - \nabla \cdot \mathbf{D}^{\alpha*} \Phi = 0 \quad \text{on } \Omega \times (0, T), \quad (25b)$$

$$\Psi = 0 \quad \text{on } \partial\Omega \times (0, T), \quad (25c)$$

$$\Psi(T) = \Theta \quad \text{on } \Omega. \quad (25d)$$

In the next lemma, an expression for the quantity  $(P_0 \varepsilon_u(T), \Theta)$  in terms of the errors  $\varepsilon'_u$ ,  $\varepsilon_q$ , the projection errors  $e_q$  and  $e'_u$ , and the solution of the dual problem will be given. In it,  $\mathbf{I}_h$  is any interpolation operator from  $L^2(\Omega)$  into  $W_h \cap H_0^1(\Omega)$ ,  $P_w$  is the  $L^2$ -projection into  $W_h$  and  $\Pi^{\text{BDM}}$  is the well-known projection associated to the lowest-order Brezzi-Douglas-Marini (BDM) space.

**Lemma 3** *Assume that  $k \geq 1$ . Then, for any  $T > 0$ ,*

$$\begin{aligned} (P_0 \varepsilon_u(T), \Theta) &= \int_0^T [(\varepsilon_q, \mathbf{D}^{\alpha*}(\nabla \mathbf{I}_h \Psi) - \Pi^{\text{BDM}} \nabla \Psi) \\ &\quad + (e_q, \mathbf{D}^{\alpha*}(\Pi^{\text{BDM}} \nabla \Psi - \nabla P_w \Psi)) + (\varepsilon'_u - e'_u, P_0 \Psi - \mathbf{I}_h \Psi)] dt. \end{aligned}$$

*Proof* Since  $\Psi(T) = \Theta$  by (25d) and  $\varepsilon_u(0) = 0$  (since  $u_h(0) = \Pi_w u_0$ ), we have

$$\begin{aligned} (P_0 \varepsilon_u(T), \Theta) &= \int_0^T [((P_0 \varepsilon_u)', \Psi) + (P_0 \varepsilon_u, \Psi')] dt \\ &= \int_0^T [(\varepsilon'_u, P_0 \Psi) + (\varepsilon_u, P_0 \nabla \cdot \mathbf{D}^{\alpha*} \Phi)] dt \end{aligned}$$

by the definition of the  $L^2$ -projection  $P_0$  and by (25b).

By the commutativity property  $P_0 \nabla \cdot = \nabla \cdot \Pi^{\text{BDM}}$  and the first error equation (11a) with  $\mathbf{r} := \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi$  (since  $k \geq 1$ ), we get for each  $t \in (0, T]$ ,

$$\begin{aligned} (\varepsilon_u, P_0 \nabla \cdot \mathbf{D}^{\alpha*} \Phi) &= (\varepsilon_u, \nabla \cdot \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi), \\ &= (\varepsilon_q, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi) + \langle \varepsilon_{\hat{u}}, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi \cdot \mathbf{n} \rangle - (e_q, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi) \\ &= (\varepsilon_q, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi) - (e_q, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi) \\ &= (\varepsilon_q, \mathbf{D}^{\alpha*}(-\Pi^{\text{BDM}} \nabla \Psi + \nabla \mathbf{I}_h \Psi)) - (\varepsilon_q, \mathbf{D}^{\alpha*}(\nabla \mathbf{I}_h \Psi)) - (e_q, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi). \end{aligned}$$

Noting that, in the second last equality we used

$$\langle \varepsilon_{\hat{u}}, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi \cdot \mathbf{n} \rangle = \langle \varepsilon_{\hat{u}}, \mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi \cdot \mathbf{n} \rangle_{\partial\Omega} = 0$$

which follows from (11d) (because  $\mathbf{D}^{\alpha*} \Pi^{\text{BDM}} \Phi \in H(\text{div}, \Omega)$ ) and the fact that  $\varepsilon_{\hat{u}} = 0$  on  $\partial\Omega$  by (11c).

But, by the error equation (11b) with  $w := \mathbf{D}^{\alpha*}(\mathbf{I}_h \Psi)$ ,

$$(\varepsilon_q, \mathbf{D}^{\alpha*}(\nabla \mathbf{I}_h \Psi)) = (\mathcal{I}^\alpha(\varepsilon'_u - e'_u), \mathbf{D}^{\alpha*}(\mathbf{I}_h \Psi)) + \langle \varepsilon_{\hat{q}} \cdot \mathbf{n}, \mathbf{D}^{\alpha*}(\mathbf{I}_h \Psi) \rangle.$$

Now, putting together all the above intermediate steps,

$$\begin{aligned} (P_0 \varepsilon_u(T), \Theta) &= \int_0^T [(\varepsilon'_u, P_0 \Psi) + (\varepsilon_{\mathbf{q}}, D^{\alpha*}(\nabla I_h \Psi) - \mathbf{\Pi}^{\text{BDM}} \nabla \Psi) \\ &\quad - (D^\alpha \mathcal{I}^\alpha(\varepsilon'_u - e'_u), I_h \Psi) - \langle \varepsilon_{\hat{\mathbf{q}}} \cdot \mathbf{n}, D^{\alpha*}(I_h \Psi) \rangle - (e_{\mathbf{q}}, D^{\alpha*} \mathbf{\Pi}^{\text{BDM}} \Phi)] dt. \end{aligned} \quad (27)$$

Changing the order of integrals and using  $\int_s^t \omega_{1-\alpha}(t-q)\omega_\alpha(q-s) dq = 1$ , we obtain the following identity:

$$\begin{aligned} D^\alpha(\mathcal{I}^\alpha(\varepsilon'_u - e'_u))(t) &= \frac{\partial}{\partial t} \int_0^t \omega_{1-\alpha}(t-q) \int_0^q \omega_\alpha(q-s)(\varepsilon_u - e_u)'(s) ds dq \\ &= \frac{\partial}{\partial t} \int_0^t (\varepsilon_u - e_u)'(s) \int_s^t \omega_{1-\alpha}(t-q)\omega_\alpha(q-s) dq ds \\ &= \frac{\partial}{\partial t} \int_0^t (\varepsilon_u - e_u)'(s) ds = (\varepsilon_u - e_u)'(t). \end{aligned}$$

In addition,  $\langle \varepsilon_{\hat{\mathbf{q}}} \cdot \mathbf{n}, D^{\alpha*}(I_h \Psi) \rangle = \langle \varepsilon_{\hat{\mathbf{q}}} \cdot \mathbf{n}, D^{\alpha*}(I_h \Psi) \rangle_{\partial\Omega} = 0$  by (11d) and the identity  $I_h \Psi = 0$  on  $\partial\Omega$  by the boundary condition of the dual problem (25c).

Hence, by using this in (27), we observe

$$\begin{aligned} (P_0 \varepsilon_u(T), \Theta) &= \int_0^T [(\varepsilon_{\mathbf{q}}, D^{\alpha*}(\nabla I_h \Psi - \mathbf{\Pi}^{\text{BDM}} \nabla \Psi)) \\ &\quad - (e_{\mathbf{q}}, D^{\alpha*} \mathbf{\Pi}^{\text{BDM}} \Phi) + (\varepsilon'_u, P_0 \Psi - I_h \Psi) + (e'_u, I_h \Psi)] dt. \end{aligned}$$

Therefore, the desired result now follows after noting that

$$- \int_0^T (e_{\mathbf{q}}, D^{\alpha*} \mathbf{\Pi}^{\text{BDM}} \Phi) dt = \int_0^T (e_{\mathbf{q}}, D^{\alpha*}(\mathbf{\Pi}^{\text{BDM}} \nabla \Psi - \nabla P_w \Psi)) dt,$$

(by (25a), the fact that  $P_w$  is the  $L^2$ -projection into  $W_h$ , and the orthogonality property of the projection  $\mathbf{\Pi}_V$ , (9a)) and that  $(e'_u, I_h \Psi) = (e'_u, I_h \Psi - P_0 \Psi)$  (by the fact that  $P_0 \Psi$  is constant on each element  $K \in \mathcal{T}_h$ , and the orthogonality property of the projection  $\Pi_w$ , (9b)). The proof is completed now.

In the next theorem we state the superconvergence estimate of the postprocessed HDG approximation. For the proof, we follow the derivation in [6, Section 5] step-by-step and use Lemma 3 instead of [6, Lemma 7].

**Theorem 3** *Assume that  $u \in \mathcal{C}^1(0, T; H^{k+2}(\Omega))$  and  $\mathbf{q} \in \mathcal{C}^1(0, T; \mathbf{H}^{k+1}(\Omega))$ . Assume also that  $\tau_K^*$  and  $1/\tau_K^{\max}$  are bounded by  $\mathbf{C}$ . Then, we have*

$$\|(u - u_h^*)(T)\| \leq C_2 \max\{1, \sqrt{\log(Th^{-2/(\alpha+1)})}\} h^{k+2} \quad \text{for } k \geq 1,$$

where the constant  $C_2$ , only depends on  $\mathbf{C}$ ,  $\alpha$ ,  $T$ ,  $\|u\|_{\mathcal{C}^1(H^{k+2})}$ , and on  $\|\mathbf{q}\|_{\mathcal{C}^1(H^{k+1})}$ .

## 5 Numerical experiments

In this section, we present numerical experiments devised to validate our theoretical predictions from HDG spatial discretizations. To do so, we use the fully discrete CN HDG scheme (8). We take the (uniform) time steps  $\delta$  to be sufficiently small so that the HDG and postprocessed HDG spatial discretizations errors are dominant. This is achieved by fixing the ratio  $\frac{\delta^2}{h^{k+2}}$  to a given number less than the unit because the time stepping CN scheme is second-order accurate provided that the exact solution is sufficiently regular.

We choose the spatial domain  $\Omega$  to be the unit interval  $(0, 1)$  and  $T = 1$  in (1). We impose homogenous Dirichlet boundary conditions and choose the source term  $f$  and the initial data  $u_0$  so that the exact solution is  $u(x, t) = t^{3-\alpha} \sin(\pi x)$ . For different values of  $\alpha$ , we obtain the history of convergence of the errors  $\|(u - u_h)(T)\|$ ,  $\|(\mathbf{q} - \mathbf{q}_h)(T)\|$  and  $\|(u - u_h^*)(T)\|$  for different values of the polynomial degree,  $k = 0, 1, 2$ . To compute the spatial  $L_2$ -norm, we apply a composite Gauss quadrature rule with 4-points on each interval of the finest spatial mesh. The numerical results (errors and convergence rates) of the experiments are presented in Tables 1 and 2. In full agreement with our theoretical results, we obtain optimal convergence rates for the HDG scheme and  $O(h^{k+2})$  superconvergence rates for the postprocessed HDG scheme.

$N$	$k = 0$					
4	5.269e-01		7.899e-01		5.048e-01	
8	3.027e-01	0.799	4.028e-01	0.972	2.922e-01	0.788
16	1.616e-01	0.905	2.025e-01	0.992	1.566e-01	0.899
32	8.342e-02	0.954	1.014e-01	0.997	8.098e-02	0.951
64	4.237e-02	0.977	5.072e-02	0.999	4.117e-02	0.976
128	2.135e-02	0.989	2.537e-02	0.999	2.076e-02	0.988
	$k = 1$					
4	6.031e-02		5.936e-02		7.401e-03	
8	1.502e-02	2.005	1.321e-02	2.165	8.835e-04	3.066
16	4.144e-03	1.858	3.487e-03	1.924	1.142e-04	2.951
32	1.048e-03	1.983	8.649e-04	2.011	1.420e-05	3.008
	$k = 2$					
4	3.960e-03		4.596e-03		8.902e-04	
8	5.059e-04	2.969	4.868e-04	3.239	5.497e-05	4.017
16	6.352e-05	2.993	5.652e-05	3.107	3.416e-06	4.008
32	7.957e-06	2.997	7.117e-06	2.989	2.153e-07	3.988

**Table 1** The errors  $\|(u_h - u)(T)\|$ ,  $\|(\mathbf{q}_h - \mathbf{q})(T)\|$  and  $\|(u_h^* - u)(T)\|$ , and the corresponding rates of convergence for  $\alpha = 0.5$  with HDG solutions of degree  $k = 0, 1, 2$ . We observe optimal convergence of order  $h^{k+1}$  for the errors in  $u_h$  and  $\mathbf{q}_h$ , and superconvergence rates of order  $h^{k+2}$  (when  $k \geq 1$ ) for the error from the postprocessed HDG solution  $u_h^*$ .

$N$	$k = 0$					
4	5.455e-01		7.705e-01		5.240e-01	
8	3.122e-01	0.805	3.088e-01	9.884	3.020e-01	0.795
16	1.661e-01	0.910	1.939e-01	1.002	1.612e-01	0.905
32	8.558e-02	0.957	9.674e-02	1.003	8.320e-02	0.954
64	4.342e-02	0.979	4.830e-02	1.002	4.225e-02	0.978
128	2.187e-02	0.989	2.413e-02	1.001	2.128e-02	0.989
	$k = 1$					
4	6.081e-02		6.005e-02		7.898e-03	
8	1.501e-02	2.018	1.321e-02	2.185	9.403e-04	3.070
16	4.154e-03	1.854	3.485e-03	1.922	1.218e-04	2.949
32	1.048e-03	1.987	8.434e-04	2.047	1.506e-05	3.015
	$k = 2$					
4	4.025e-03		4.978e-03		1.079e-03	
8	5.088e-04	2.984	5.014e-04	3.312	6.698e-05	4.010
16	6.367e-05	2.998	5.701e-05	3.137	4.167e-06	4.007
32	7.892e-06	3.012	7.031e-06	3.019	2.641e-07	3.980

**Table 2** The errors  $\|(u_h - u)(T)\|$ ,  $\|(\mathbf{q}_h - \mathbf{q})(T)\|$  and  $\|(u_h^* - u)(T)\|$ , and the corresponding rates of convergence for  $\alpha = 0.7$  with HDG solutions of degree  $k = 0, 1, 2$ .

## References

1. B. Chabaud and B. Cockburn, Uniform-in-time superconvergence of HDG methods for the heat equation, *Math. Comp.*, 81, (2012) 107–129.
2. C-M. Chen, F. Liu, V. Anh and I. Turner, Numerical methods for solving a two-dimensional variable-order anomalous sub-diffusion equation, *Math. Comp.*, 81, (2012) 345–366.
3. J. Chen, F. Liu, Q. Liu, X. Chen, V. Anh, I. Turner, K. Burrage, Numerical simulation for the three-dimension fractional sub-diffusion equation, *Appl. Math. Model.*, 38, (2014) 3695–3705.
4. B. Cockburn, J. Gopalakrishnan and R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems, *SIAM J. Numer. Anal.*, 47, (2009) 1319–1365.
5. B. Cockburn, J. Gopalakrishnan and F.-J. Sayas, A projection-based error analysis of HDG methods, *Math. Comp.*, 79, (2010) 1351–1367.
6. B. Cockburn and K. Mustapha, A hybridizable discontinuous Galerkin method for fractional diffusion problems, *Numer. Math.*, (2014), DOI 10.1007/s00211-014-0661-x.
7. B. Cockburn, W. Qiu and K. Shi, Conditions for superconvergence of HDG methods for second-order elliptic problems, *Math. Comp.*, 81, (2012) 1327–1353.
8. E. Cuesta, C. Lubich and C. Palencia, Convolution quadrature time discretization of fractional diffusive-wave equations, *Math. Comp.*, 75, (2006) 673–696.
9. M. Cui, Compact finite difference method for the fractional diffusion equation, *J. Comput. Phys.*, 228, (2009) 7792–7804.
10. M. Cui, Convergence analysis of high-order compact alternating direction implicit schemes for the two-dimensional time fractional diffusion equation, *Numer. Algor.*, 62, (2013) 383–409.
11. G.G. Gao and Z.Z. Sun, A box-type scheme for fractional sub-diffusion equation with Neumann boundary conditions, *J. Comput. Phys.*, 230, (2011) 6061–6074.
12. L. Gastaldi and R.H. Nochetto, Sharp maximum norm error estimates for general mixed finite element approximations to second order elliptic equations, *RAIRO Modél. Math. Anal. Numér.*, 23, (1989) 103–128.
13. B. Jin, R. Lazarov, and Z. Zhou, On two schemes for fractional diffusion and diffusion-wave equation, <http://arxiv.org/pdf/1404.3800v1.pdf>, (2014).

14. A.A. Kilbas, H.M. Srivastava and J.J. Trujillo, Theory and Applications of Fractional Differential Equations, Volume 204 (North-Holland Mathematics Studies), 2006.
15. R.M. Kirby, S.J. Sherwin and B. Cockburn, To HDG or to CG: A comparative study, *J. Sci. Comput.*, 51, (2012) 183–212.
16. X. Li and C. Xu, A space-time spectral method for the time fractional diffusion equation, *SIAM J. Numer. Anal.*, 47, (2009), 2108–2131.
17. Y. Lin and C. Xu, Finite difference/spectral approximations for the time-fractional diffusion equation, *J. Comput. Phys.*, 225, (2007), 1533–1552.
18. W. McLean, Regularity of solutions to a time-fractional diffusion equation, *ANZIAM J.*, 52, (2010) 123–138.
19. W. McLean and K. Mustapha, Convergence analysis of a discontinuous Galerkin method for a sub-diffusion equation, *Numer. Algor.*, 52, (2009) 69–88.
20. W. McLean and K. Mustapha, Error analysis of a discontinuous Galerkin method for a fractional diffusion equation with a non-smooth initial data, *J. Comput. Phys.*, (2014), DOI: 10.1016/j.jcp.2014.08.050.
21. R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, *Physics Reports* 339, (2000) 1–77.
22. K. Mustapha, An implicit finite difference time-stepping method for a sub-diffusion equation, with spatial discretization by finite elements, *IMA J. Numer. Anal.*, 31, (2011) 719–739.
23. K. Mustapha, Time-stepping discontinuous Galerkin methods for fractional diffusion problems, *Numer. Math.*, (2014), DOI 10.1007/s00211-014-0669-2.
24. K. Mustapha and J. AlMutawa, A finite difference method for an anomalous sub-diffusion equation, theory and applications, *Numer. Algor.*, 61, (2012) 525–543.
25. K. Mustapha, B. Abdallah and K. Furati, A discontinuous Petrov-Galerkin method for time-fractional diffusion equations, *SIAM J. Numer. Anal.*, 52 (2014) 2512–2529.
26. K. Mustapha and W. McLean, Piecewise-linear, discontinuous Galerkin method for a fractional diffusion equation, *Numer. Algor.*, 56, (2011) 159–184.
27. K. Mustapha and W. McLean, Uniform convergence for a discontinuous Galerkin, time stepping method applied to a fractional diffusion equation, *IMA J. Numer. Anal.*, 32, (2012) 906–925.
28. K. Mustapha and W. McLean, Superconvergence of a discontinuous Galerkin method for fractional diffusion and wave equations, *SIAM J. Numer. Anal.*, 51, (2013) 491–515.
29. K. Mustapha and D. Schötzau, Well-posedness of  $hp$ -version discontinuous Galerkin methods for fractional diffusion wave equations, *IMA J. Numer. Anal.*, 34, (2014) 1426–1446.
30. N.C. Nguyen, J. Peraire and B. Cockburn, Hybridizable discontinuous Galerkin methods, *Proceedings of the International Conference on Spectral and High Order Methods* (Trondheim, Norway), Lect. Notes Comput. Sci. Engrg., Springer Verlag, June 2009.
31. J. Quintana-Murillo and S.B. Yuste, An explicit difference method for solving fractional diffusion and diffusion-wave equations in the Caputo form, *J. Comput. Nonlin. Dyn.*, 6, (2011) 021014.
32. N. H. Sweilam, M. M. Khader and A. M. S. Mahdy, Crank-Nicolson finite difference method for solving time-fractional diffusion equation, *J. Fract. Cal. Appl.*, 2 (2012) 1–9.
33. W. Wyss, Fractional diffusion equation, *J. Math. Phys.*, 27, (1986) 2782–2785.
34. Q. Xu and Z. Zheng, Discontinuous Galerkin method for time fractional diffusion equation, *J. Informat. Comput. Sci.*, 10 (2013) 3253–3264.
35. S.B. Yuste and J. Quintana-Murillo, On Three Explicit Difference Schemes for Fractional Diffusion and Diffusion-Wave Equations, *Phys. Scripta T136*, (2009) 014025.
36. F. Zengo, C. LI, F. Liu, I. Turner, The use of finite difference/element approaches for solving the time-fractional subdiffusion equation, *SIAM J. Sci. Comput.*, 35, (2013) A2976–A3000.
37. Y.-N. Zhang and Z.-Z. Sun, Alternating direction implicit schemes for the two-dimensional fractional sub-diffusion equation, *J. Comput. Phys.*, 230, (2011) 8713–8728.
38. X. Zhao and Z.-z. Sun, A box-type scheme for fractional sub-diffusion equation with Neumann boundary conditions, *J. Comput. Phys.*, 230, (2011) 6061–6074.
39. X. Zhao and Q. Xu, Efficient numerical schemes for fractional sub-diffusion equation with the spatially variable coefficient, *Appl. Math. Model.*, 38, (2014) 3848–3859.